



The Interpretation of the Laakso-Taagepera Effective Number of Parties

Jean-François Caulier

► To cite this version:

Jean-François Caulier. The Interpretation of the Laakso-Taagepera Effective Number of Parties. 2011. halshs-00565315

HAL Id: halshs-00565315

<https://shs.hal.science/halshs-00565315>

Submitted on 11 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**The Interpretation of the Laakso-Taagepera Effective
Number of Parties**

Jean-François CAULIER

2011.06



The Interpretation of the Laakso-Taagepera Effective Number of Parties*

Jean-François Caulier[†]
Centre d'Economie de la Sorbonne, University of Paris I
106-112, Bd de l'Hôpital
75013 Paris, France

February 3, 2011

Abstract

In this paper, we present a general statistical framework within which we can draw a new interpretation of the Laakso-Taagepera effective number of parties fragmentation index. With the particular method of sampling with probability proportional to the party sizes, we show that the Laakso-Taagepera effective number of parties is the inverse of the size biased version of the traditional expected party size in shares. Further, we provide an axiomatic definition of the Laakso-Taagepera effective number of parties.

Keywords: Fragmentation, effective number of parties, concentration index, size biased sampling, length biased sampling

Résumé

Nous présentons dans cet article un environnement statistique général dans lequel nous pouvons donner une interprétation nouvelle du nombre effectif de partis de Laakso-Taagepera comme indice de fragmentation. Grâce à la méthode particulière d'échantillonnage dans laquelle les probabilités sont proportionnelles aux tailles des partis, nous montrons que le nombre effectif de partis de Laakso-Taagepera n'est rien d'autre que la version biaisée par la taille de l'espérance normale de la taille des partis. De plus, nous fournissons une définition axiomatique du nombre effectif de partis de Laakso-Taagepera.

JEL classification: C65, D71, D72

Mot-clés : Fragmentation, nombre effectif de partis, indice de concentration, échantillonnage biaisée par la taille

*This working paper is a preliminary version of a joint article with Patrick Dumont (Université du Luxembourg).

[†]The author thanks Jean-Luc Marichal for the help on the axiomatic part of the paper. We also thank Agnieszka Rusinowska and the participants to the journées doctorales de l'ADRES, Paris January 21-22nd 2011 for helpful comments.

Introduction

What do we mean by the *fragmentation* of a party-system ? What is the definition of such a widely used concept in the political science field ? Why is it important to assess the degree of fragmentation of party-systems ? More importantly why the concept receives hearty agreement for its use when there is no unanimous consensus in its meaning ? Our aim is to discover or understand the definition that has to be given to the notion of *fragmentation* on the basis of its most accepted operationalization : the Effective Number of Parties (ENP). Our task is motivated by the work of Feld and Grofman [2007] that present an alternative way to compute the Effective Number of Parties in terms of well-known statistics (mean and variance). Their work used as a springboard for ours, we seek to dive more deeper in the mathematical properties of the ENP in order to understand exactly its exact meaning and interpretation.

An important aspect of classification of party-systems is the degree of competition among political parties which in turn, depends *ceteris paribus* on the number and sizes of the elected parties. A system is said to be *fragmented* if it is composed of many elected parties. Hence, at a first approximation we can think that a key feature for a definition of the notion of *fragmentation* refers to the number of elected parties (see Elkins [1974],p.683). Of course the number of parties alone does not tell anything on the size distribution of seats and is of poor information. Some authors have proposed to enrich the concept of fragmentation in order to go beyond the only consideration of the number of parties and the distribution of seats, such as in Blau [2008] or Ricciuti [2004]. Despite these efforts, the quantitative very aspect of fragmentation remains deeply rooted in the field. Since the first proposal for classification in term of two-party and multi-party proposed by Duverger (Duverger [1959]), scholars have constantly tried to give more precise quantifications of the number of competing parties and distribution of sizes. One such measure has reached a high degree of consensus among scholars : the Laakso-Taagepera Effective Number of Parties (ENP) (see Lijphart [1994] p.68 and Laakso and Taagepera [1979]). Despite the drawbacks and flaws of the ENP, stressed by some authors suggesting new measures (see Dumont and Caulier [2003], Dunleavy and Boucek [2003], Golosov [2010], Kline [2009], Molinar [1991]), the ENP remains the most constant used, if not the only one, measure of party-system fragmentation. It thus seems that if we are in state to correctly and consistently give a correct interpretation to the ENP, we will directly get a definition for the concept of fragmentation, given that the ENP is “The” measure of fragmentation.

The ENP in a mean and variance framework

We introduce some notation in order to present the classical formula for the ENP and the proposed reformulation of Feld and Grofman [2007] in terms of mean and variance.

Let N denote a party-system composed of n political parties, n a finite integer. Each party i owns $x_i \in \mathbb{R}_+$ seats in the party-system, with \mathbb{R}_+ the set of nonnegative real numbers and $\bar{x} = \sum_{i=1}^n x_i$ is the total number of seats in the system. We assume that \bar{x} is known and fixed for a given party-system. We denote by $s_i = x_i/\bar{x}$ the shares of seats of party i , for each i in N . A party-system N consists either in a distribution of seats $\mathbf{x} \in \mathbb{R}_+^n$ or shares $\mathbf{s} \in [0, 1]^n$ with \mathbb{R}_+^n and $[0, 1]^n$ the n -fold cartesian product of \mathbb{R}_+ and $[0, 1]$ respectively. The *Effective Number of Parties* (Laakso and Taagepera [1979]) of a party-system N is a function $\text{ENP} : \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ defined by

$$\text{ENP}(\mathbf{x}) = \frac{1}{\sum_{i=1}^n (x_i/\bar{x})^2} = \left(\sum_{i=1}^n s_i^2 \right)^{-1}$$

the inverse of the sum of squared shares.

Some basic statistics : For a distribution of seats \mathbf{x} , let

$$\mu(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n x_i \tag{1}$$

the mean number of seats among parties,

$$\sigma^2(\mathbf{x}) = 1/n \sum_{i=1}^n (x_i - \mu(\mathbf{x}))^2 \tag{2}$$

the variance of \mathbf{x} . Correspondingly we have $\mu(\mathbf{s})$ and $\sigma^2(\mathbf{s})$ the mean and variance of \mathbf{s} . Note that $\mu(\mathbf{s}) = 1/n$. Alternatively we can calculate the variance by

$$\sigma^2(\mathbf{s}) = \mu(\mathbf{s}^2) - \mu^2(\mathbf{s}) = \frac{1}{n} \sum_{i=1}^n s_i^2 - \frac{1}{n^2},$$

that can be rewritten

$$\sum_{i=1}^n s_i^2 = n \left(\sigma^2(\mathbf{s}) + \frac{1}{n^2} \right) = n\sigma^2(\mathbf{s}) + \mu(\mathbf{s})$$

establishing the identity

$$\text{ENP}(\mathbf{x}) = n/(1 + n^2\sigma^2(\mathbf{s})) = 1/(\mu(\mathbf{s}) + n\sigma^2(\mathbf{s})) \tag{3}$$

as in Feld and Grofman [2007] (equation (10), p.105).¹

¹The same derivation can be obtained from the relation $\text{ENP}^{-1} = \frac{1}{n} + n\sigma^2(\mathbf{s})$ in Golosov [2010] (equation (1) p.2).

If relation 3 is interesting as it allows the calculation of ENP using two well-known statistics, the mean and the variance, or even simpler only by knowing n and the variance (since the mean of n shares is always $1/n$), it could then mislead to the conclusion that the ENP is merely a reformulation of the variance of a distribution of seat shares. Still, we must bare in mind that the reformulation of an expression is not always insignificant and indeed, this reformulation of ENP can shed the light on two possible paths of interpretations. The first concerns the relation between ENP and concentration indices and the second one concerns the statistical interpretation of the ENP. Both kinds of interpretations explain why the ENP can be expressed only as a function of n and the variance.

The ENP as number equivalent index

One of the reasons explaining the success story of the ENP can be found on its intuitive meaning : “The effective number of components is the easiest to visualize in concrete terms: $[ENP]=2.28$ directly tells us that there are more than two but definitely less than three major parties [...] One can ask uninitiated students to estimate the effective number of parties and they respond with values approximating $[ENP]$ ” (in Taagepera and Shugart [1989], p.80).

The ENP is a *number equivalent index*. This type of index requires that if all the parties in the party-system have the same number of seats, the value taken by the index is the actual number of parties (see Adelman [1969]). The difficulty in assessing the fragmentation of a party-system lies in the variability among parties sizes. When all the parties have an equal number of seats, there is no inequality in the distribution of seats, the actual number of parties does convey some sense. If a number equivalent index evaluated for a party-system achieves a value of k , we can say that the party-system is as fragmented as if there were k parties of equal size. Of course a number equivalent index does not always achieve integer values but this must not create problems, first because it then tells us where the party-systems stands according to situations with clearcut meaning (if the index is 1.5, you can say that it is half fragmented than 3 or that it is just the halfway between 1 and 2 equal parties), secondly no one seems to be offended by an assertion stating that last year, each american woman has given birth to one and half children.

According to Blackorby et al. [1982], number equivalent indexes are inverse of concentration measures. Indeed, the link between the ENP and the Hirschman-Herfindahl concentration index is well known (see Laakso and Taagepera [1979] or Feld and Grofman [2007]). The ENP is simply the inverse of the Hirschman-Herfindahl (HH). Applied to a party-system, the HH is calculated as $HH(\mathbf{s}) = 1/ENP(\mathbf{s}) = \sum_{i=1}^n s_i^2$. The HH index achieves (the-

oretically) values between 0 and 1 indicating the degree of concentration. An increase in the value of a number equivalent corresponds to a reduction in the degree of concentration. A party-system highly concentrated ($HH \rightarrow 1$) displays most probably a dominant party owning most of the available seats and is thus poorly fragmented, with a low ENP close to 1 (the lower bound for the ENP).

Concentration indices are not inequality indices, even though both are related. Measures of concentration deal with the presence of dominant elements in term of sizes in the distribution under study. Inequality indices are concerned with the presence of small size elements by comparison of large size elements in the distribution. To stress the distinction between the aspects tackled by the measures, imagine a distribution with a giant size element. If some tiny size elements enter into the distribution, a concentration index won't be affected whereas the inequality will greatly increase. The variance of a distribution is not an adequate inequality index. If we double each entry in a given distribution such that the mean is double and the shape of the distribution remains unchanged, the variance would quadruple, while common sense would difficultly accept that inequality is aggravated. One way to transform the variance in an inequality index is to standardize it, obtaining the *coefficient of variation* :

$$c(\mathbf{x}) = \frac{\sqrt{\sigma^2(\mathbf{x})}}{\mu(\mathbf{x})}$$

or for $\mathbf{s} = \frac{\mathbf{x}}{\bar{x}}$,

$$c(\mathbf{s}) = n\sqrt{\sigma^2(\mathbf{s})},$$

a function of n and $\sigma^2(\mathbf{s})$. The coefficient of variation records the relative dispersion of distributions and is particularly adequate to compare distributions with different means.

The precise link between concentration and inequality indices has been demonstrated first in Davies [1979] and then by Blackorby et al. [1982]. They show that the inverse of a concentration measure (a number equivalent index) can be written as a function of the number of elements n and a measure of inequality (the coefficient of variation for example). Accordingly, the ENP can be written as a function $f(n, c(\mathbf{s}))$ where f is increasing in n and decreasing in $c(\mathbf{s})$. Let us rewrite equation 3 as

$$\text{ENP}(\mathbf{s}) = \frac{n}{1 + c^2(\mathbf{s})} = \frac{n}{1 + \left(\frac{\sqrt{\sigma^2(\mathbf{s})}}{\mu(\mathbf{s})}\right)^2} = \frac{n}{1 + n^2\sigma^2(\mathbf{s})}$$

the desired relationship. Holding $c(\mathbf{s})$ fixed, we see that ENP is increasing in n and decreasing in $c(\mathbf{s})$ for a fixed n . The ENP being a number equivalent index, given the existing results in the literature, it is not a surprise that the ENP can be expressed as a function of n and σ^2 .

The ENP as statistic

In this section we present the second line of interpretation of the ENP in term of statistic. We build on this interpretation on the story told by Feld and Grofman [2007], the so-called “class-size paradox”², that we place in a more general context. As Feld and Grofman [2007] state in footnote 3, p.103, this paradox is a well-known phenomenon in the statistics literature, which can be found under various names whose most frequent are “methods of ascertainment” (Fischer [1934]), “length-biased sampling” (Cox and Lewis [1966], Zelen [1972]), “the waiting time paradox” (Feller [1971], p.12), “size-biased sampling” (Scheaffer [1972]), “the inspection paradox” (Ross [1972]), “weighted distributions” (Patil et al. [1988]), ...

All these papers deal with the same problem of sampling bias. In the class-size paradox, you want to estimate the average size of classes in a school by asking students the size of their class. This method introduces an estimation bias since you are sampling students and not classes. Suppose that you repeatedly go for a walk after dinner and meet ten people in your neighborhood out of which seven have a dog, you won’t surely conclude that 70% of your neighbors have a dog. When you arrive randomly at a bus stop, due to the variation of circulation of buses, you will have to wait more than half the average time between two buses. You are more likely to arrive during longer intervals between two buses and must wait on average more than the average inter-arrival time. A radar fixed at an arbitrary location on a highway has more chances to record faster cars, slow-developing tumors have more chances to be detected in random medical screenings of a population, ... In all these examples, the variation in sizes or length matters a lot, and the problem disappears as soon as there is no variation in size.

We have already seen that there is no difficulty to assess the fragmentation in a party-system with equal size parties, but the raw number of parties is a useless information as soon as there is variability in size. The intuitive explanation usually put forward for the ENP is that some weighting is necessary. Bigger parties receive more weight in the computation, and to avoid arbitrary weights, we use weight proportional to size (Taagepera [2007] p.58). This amounts to introducing a sampling bias. We now present the ENP formula in the context of sampling bias to show the logic of its weighting and to present the consequent interpretation.

We start with a finite party-system N with n parties. Let $i = 1, \dots, n$ denote the labels of parties in N . The distribution of seats among parties in N is $\mathbf{x} = x_1, \dots, x_n$ with x_i being the number of seats own by party i . We design a random experiment consisting in drawing a party i . The discrete random variable X is the number of seats own by party i . The probability function associated to X is $p^X(x_i) = P(X = x_i)$. The expected value of X ,

²(see also Feld [1991], Feld and Grofman [1977], Feld and Grofman [2010])

denoted by $E[X]$, is defined by

$$E[X] = \sum_{i=1}^n x_i p^X(x_i).$$

This expression is the expected number of seats given the random selection process of parties. Note that $E[X]$ corresponds to the mean number of seats μ in equation (1) if and only if $p^X(x_i) = \frac{1}{n}$ for all $i \in N$, when all parties have the same probability to be selected. The variance of X , denoted by $\text{Var}(X)$ is defined by

$$\text{Var}(X) = E[(X - E[X])^2] = \sum_{i=1}^n p^X(x_i)(x_i - E[X])^2 = E[X^2] - E^2[X].$$

Again, the variance of X corresponds to the variance of the distribution σ^2 in equation (2) if and only if $p^X(x_i) = \frac{1}{n}$ for all $i \in N$.

We now design another experiment that consists in drawing a seat s at random, with s the label of the seat, taking values from 1 to \bar{x} , the total number of seats. To each seat corresponds one and only one party. The discrete random variable X^* is the number of seats of the corresponding party owning s , i.e. if s belonging to party j has been chosen by the random experiment, then $X^* = x_j$. The probability function associated to X^* is $p^{X^*}(x_i) = P(X^* = x_i)$. This experiment is an indirect way to choose a party.³ Correspondingly, the expected value of X^* is defined by

$$E[X^*] = \sum_{i=1}^n x_i p^{X^*}(x_i).$$

and its variance by

$$E[(X^* - E[X^*])^2] = \sum_{i=1}^n p^{X^*}(x_i)(x_i - E[X^*])^2 = E[X^{*2}] - E^2[X^*].$$

According to the random experiment chosen, there exist two different ways to estimate the mean number of seats by party. In the second experiment, selecting a seat is an indirect way of picking a party. Given the number of seats owned by parties, parties are selected with a probability proportional to the number of seats they have. The second experiment is a method of sampling with bias proportional to size. X^* is called the *size-biased version* of X and the probability function of X^* is called the *size-biased probability function* of X

$$p^{X^*}(X^*) = \frac{X p^X(X)}{E[X]}$$

³Note how our constructions mimic the procedure of computing the averages of class size in Feld and Grofman [2007].

and thus

$$E[X^*] = \sum_{i=1}^n x_i p^{X^*}(x_i) = \frac{\sum_{i=1}^n x_i^2 p^X(x_i)}{E[X]} = \frac{\text{Var}(X) + E^2[X]}{E[X]}$$

that can further be simplified

$$E[X] + \frac{\text{Var}(X)}{E[X]} = E[X^*]. \quad (4)$$

The difference $E[X^*] - E[X] = \frac{\text{Var}(X)}{E[X]}$ is the bias involved estimating $E[X]$ by $E[X^*]$.

Using this general setting, we are in state to give the interpretation of the ENP in terms of the two random processes we have just designed.

We recall that the population of \bar{x} seats is distributed across n parties and we denote x_i the number of seats of party i , $i = 1, \dots, n$. Note that $\sum_{i=1}^n x_i = \bar{x}$. Then the expected number of seats in a equally randomly selected party, i.e. the average party size, is given by equation (1)

$$\mu(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n x_i = \frac{\bar{x}}{n}$$

with the corresponding variance $\sigma^2(\mathbf{x})$ (equation (2)). Now if $\bar{x} = 1$ we get by equation (1)

$$\mu(\mathbf{s}) = \frac{1}{n}.$$

With the second random process consisting in drawing a seat at random, we get the expected size (in number of seats) of a party containing an equally randomly selected seat and it is given by

$$\mu^*(\mathbf{x}) = \sum_{i=1}^n \frac{x_i}{\bar{x}} x_i = \sum_{i=1}^n \frac{x_i^2}{\bar{x}} = \frac{x_1^2 + \dots + x_n^2}{x_1 + \dots + x_n} \quad (5)$$

that comes from the reasoning of choosing a seat with probability $\frac{1}{\bar{x}}$ and recording the size of the party owning the seat, so that if you record the score for all seats, out of \bar{x} seats, you will record x_i times the size x_i for each possible $i = 1, \dots, n$. For each value x_i in the first distribution, you get x_i^2 entries in the second distribution. If $\bar{x} = 1$ we get by equation (5)

$$\mu^*(\mathbf{s}) = \sum_{i=1}^n s_i^2 = \frac{1}{\text{ENP}(\mathbf{x})}. \quad (6)$$

Using equation (4) we have

$$\mu^*(\mathbf{s}) = \mu(\mathbf{s}) + \frac{\sigma^2(\mathbf{s})}{\mu(\mathbf{s})} = \mu(\mathbf{s}) + n\sigma^2(\mathbf{s}). \quad (7)$$

Combining equations (6) and (7), we get the identity in equation (3) of Feld and Grofman [2007] (equation (10), p.105).

For example, suppose two parties with $s_1 = 0.6$ and $s_2 = 0.4$, draw a seat at random and record the size of its party, with 60% of chances you will record 0.6 and with 40% of chances you will record 0.4, making the expected size of the party for the average seat $60\%(0.6) + 40\%(0.4) = 0.52$, contrasting with the average size of parties of $\frac{0.6+0.4}{2} = 0.5$. The expected size from drawing a party at random is 0.5 whereas the expected size of a party from drawing at random a seat is 0.52. Both means can be obtained by direct calculation using equations (1) and (5) or by equation (7) : $\mu^*(0.6, 0.4) = 0.5 + \frac{0.01}{0.5} = 0.52$

In term of interpretation, $\mu(\mathbf{x})$ is the average size for a party, or the size of *the average party*, whereas $\mu^*(\mathbf{x})$ is the size of the party owning *the average seat*. Correspondingly, $\mu(\mathbf{s})$ are the seat shares of the average party and $\mu^*(\mathbf{s})$ are the seat shares of the party owning the average seat. The difference between $\mu(\mathbf{s})$ and $\mu^*(\mathbf{s})$ is the *selection bias* due to the method of sampling. Here, we consider that the population elements are the seats whereas the sampling or observational units are the parties, considered as groups of seats. Hence, $\mu(\mathbf{s})$ is the *group mean* and $\mu^*(\mathbf{s})$ is the *element mean*. By equation (7) we see that the element mean is always greater than the group mean unless there is no variability (measured by $\sigma^2(\mathbf{s})$) in the distribution of seat shares. And this is exactly the rationale behind the ENP : we want to attach more importance to bigger parties. If you select a party indirectly by picking a seat, you have more chances to choose a bigger party. The bigger the party, the more chances the seat chosen at random belongs to the party. On the contrary, the probability that the randomly chosen seat belongs to a smaller party is smaller. This process of random group selection is known in the literature as *selection of groups with probability proportional to size* (see Kish [1965]).

This leads to a new interpretation of the ENP. The inverse of the ENP is the average size (in seat shares) of the party to which belongs a randomly chosen seat denoted $\mu^*(\mathbf{s})$. Pick at random a seat, it will belong to a party with $1/\text{ENP}$ seat shares. Generally speaking, an average is simply a total number of units divided by the number of elements. Normalizing units in seat shares, we get 1 as the total number of seat shares. The number of elements in this line of interpretation is measured by an *effective* number of parties. Hence we can call $1/\text{ENP} = \mu^*(\mathbf{s})$ the *effective average size* in seat shares.⁴ Again, if there is no variability in size among parties, the ENP is the actual number of parties n , in that case, the effective average size $\mu^*(\mathbf{s})$ is simply $1/n$, the average size in seat shares $\mu(\mathbf{s})$. In a given assembly, we

⁴See Hannah and Kay [1977].

tautologically always have $\sum_{i=1}^n s_i = 1$, then

$$\frac{1}{\text{ENP}} = \mu^*(\mathbf{s}) \Leftrightarrow \text{ENP} = \frac{1}{\mu^*(\mathbf{s})}$$

the effective number of parties is the total number of seat shares divided by the effective average size.

Definition. Let $\mathbf{x} = x_1, \dots, x_n$ a distribution of seats among the n political parties in a finite party-system N and $\bar{x} = \sum_{i=1}^n x_i$ the total number of seats in N . The corresponding distribution of seat shares is denoted $\mathbf{s} = s_1, \dots, s_n$ with $s_i = \frac{x_i}{\bar{x}}$ for $i = 1, \dots, n$.

- Draw a seat at random with probability $\frac{1}{\bar{x}}$ and record the size of the corresponding party, then $\mu^*(\mathbf{x}) = \frac{x_1^2 + \dots + x_n^2}{x_1 + \dots + x_n}$ is the expected size of the party owning the seat, or the **effective average size**.
- Correspondingly, $\mu^*(\mathbf{s}) = s_1^2 + \dots + s_n^2$ is the expected shares of the party owning the seat, or the **Hirschman-Herfindahl effective average seat shares**.
- With the same random process, the **Laakso-Taagepera effective number of parties** $\text{ENP}(\mathbf{x}) = \frac{1}{\mu^*(\mathbf{s})}$ is the expected number of parties having the same size of the party owning the randomly selected seat.

In light of this framework, the new interpretation of the ENP is quite interesting. If you repeatedly draw a seat at random, then you know that in average, there will be $\text{ENP}(\mathbf{x})$ parties of the same size of the party from which the chosen seat has been drawn, including this party. We also provide a new interpretation of the classical concentration Hirschman-Herfindahl index as the effective average size in seat shares, that contrasts with the usual interpretation as the probability that two deputies picked at random from among the parties will be of the same party (see Golosov [2010], Dalton [2008]).⁵

The contraharmonic mean

In this section we briefly comment another occurrence of $\mu^*(\mathbf{x}) = \frac{x_1^2 + \dots + x_n^2}{x_1 + \dots + x_n}$ in the literature as *contraharmonic mean* or *antiharmonic mean*.

the first reason we can put forward in order to explain the given name can be found for $n = 2$ (see Lann and Falk [2005]), where its relationship

⁵Draw a deputy at random, with probability $1/s_i$ he will belong to party i , draw another deputy, with the same probability he will belong to party i , so that with probability $1/s_i^2$ both will belong to the same party.

with respect to the harmonic mean is plain. The arithmetic mean of two numbers a and b is

$$A = \frac{a+b}{2}.$$

The harmonic mean of two numbers a and b is

$$H = \frac{2ab}{a+b}.$$

The contraharmonic mean of two numbers a and b is

$$C = \frac{a^2 + b^2}{a+b}.$$

In these particular cases, C and H are equally distant from A , i.e.

$$C - A = A - H = \frac{(a-b)^2}{2(a+b)}.$$

Relationship that can be verified with $C = 2A - H$. The distance between C and A being the same as the one between A and H but on the other way can be considered as a good justification for the name contra or antiharmonic.

Alas, this relationship only holds for $n = 2$. Nevertheless, there still exists a relationship between the contraharmonic and harmonic means. Let X be a random variable and X^* its size-biased version (see previous section or appendix). Then it can be shown that (see Stein and Dattero [1985]) :

$$E \left[\frac{1}{X^*} \right] = \frac{1}{E[X]}.$$

Hence, in calculating the expectation of the inverse of the random variables X^* , i.e. the harmonic mean of X^* , we get the inverse of the expectation of X . Hence, in order to estimate the expectation of X , it is sufficient to take the *inverse of the harmonic mean* of X^* . This reason can constitute another argument for the given name of contraharmonic.

Axiomatic characterization of the ENP

In this section, we present a set of independent postulates or axioms that, if combined, fully characterize the ENP and its inverse, the HH. Note that another axiomatization of the Hirschman-Herfindahl index already exists in the literature (see Chakravarty and Eichorn [1991]) that completely differs from ours.⁶ We adopt the usual notation : $\mathbf{x} \in \mathbb{R}_+^n$ is the distribution of

⁶In Chakravarty and Eichorn [1991], the authors characterize a general class of concentration indices, known as the Hannah-Kay class (see Hannah and Kay [1977]) to which the Hirschman-Herfindahl index is a particular case. They show that among the self-weighted quasi-linear means, the Hannah-Kay family is the only one to satisfy a replication axiom.

seats among the n parties, for any $n \in \mathbb{N}$. The corresponding shares is denoted $\mathbf{s} \in [0, 1]^n$, $s_i = x_i/\bar{x}$ with $\bar{x} = \sum_{i=1}^n x_i$. The Laakso-Taagepera Effective Number of Parties is a function $ENP : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ such that

$$ENP(\mathbf{x}) = \left(\sum_{i=1}^n s_i^2 \right)^{-1}$$

and the Hirschman-Herfindahl concentration index is a function $HH : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ such that

$$HH(\mathbf{x}) = 1/ENP(\mathbf{x}) = \sum_{i=1}^n s_i^2.$$

We first provide the characterization of HH as measure of concentration.

Axiom 1 (Homogeneity of degree zero). *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, for all scalars $k > 0$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is homogeneous of degree zero if and only if*

$$C(\mathbf{x}) = C(k\mathbf{x})$$

All what matters in the measurement of concentration are shares and not absolute levels. An index of concentration is independent of the unit in which the components of \mathbf{x} are measured. In particular, measuring the concentration on a given distribution or on the distribution expressed in percentages should not change the degree of concentration, making concentration a relative index. We can thus state the axiom in a weaker form :

Axiom 2 (Relative index). *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is a relative index if and only if*

$$C(\mathbf{x}) = C\left(\frac{\mathbf{x}}{\bar{x}}\right).$$

Thus, if an index C is homogeneous of degree zero, C can be expressed as a relative index.

Axiom 3 (Reflexivity). *For all $n \in \mathbb{N}$, for all $a \in \mathbb{R}_+$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is reflexive if and only if*

$$C(\mathbf{1}a) = 1/n$$

with $\mathbf{1} \in \mathbb{R}_+^n = 1, \dots, 1$.

That is, if the n quantities are equal, then the value taken by function C as effective average size is simply $1/n$. In particular, if there is only one quantity in a distribution $\mathbf{x} \in \mathbb{R}^+ = \{x\}$, then $C(x) = 1$.

Axiom 4 (Recursivity). *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is recursive if and only if*

$$C(x_1, x_2, x_3, \dots, x_n) = C(x_1 + x_2, x_3, \dots, x_n) - 2 \frac{x_1 x_2}{\bar{x}^2}.$$

This axiom belongs to the general class of recursivity properties

$$C(x_1, x_2, x_3, \dots, x_n) = C(x_1 + x_2, x_3, \dots, x_n) + G(x_1, x_2)$$

for functions $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ and $G : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ where different generating functions G have been given in the literature in order to characterize information measures such as entropy (see Ebanks et al. [1998], chap. 3 or Aczel and Daroczy [1975]). The motivation for this recursivity axiom is to compare the concentration of two distributions with the same aggregate quantities and one distribution has been obtained from the other by merging in one entry the first two entries of the original distribution. In that case, by how much the concentration increases ? The difference in term of concentration when x_1 and x_2 from a distribution $\mathbf{x} \in \mathbb{R}_+^n$ merge in $x_1 + x_2$ in a distribution $\mathbf{x}' \in \mathbb{R}_+^{n-1}$ is given by

$$C(\mathbf{x}') - C(\mathbf{x}) = G(x_1, x_2) = -\frac{2}{\bar{x}} x_1 x_2$$

that can be rewritten

$$\frac{C(\mathbf{x}') - C(\mathbf{x})}{\frac{x_2}{\bar{x}}} = -2 \frac{x_1}{\bar{x}} \quad (8)$$

that is, the average concentration increase in term of the shares of x_2 is a linear function $g(x_1/\bar{x}) = 2x_1/\bar{x}$. Since we can write equation 8 as

$$\frac{C(\mathbf{x}') - C(\mathbf{x})}{\frac{x_1}{\bar{x}}} = -2 \frac{x_2}{\bar{x}}$$

we have

$$G(x_1, x_2) = g(x_1/\bar{x})x_2 = g(x_2/\bar{x})x_1 \Rightarrow g(x) = kx$$

with $x \in \mathbb{R}_+$ and k an arbitrary constant, in particular in this case $k = -2/\bar{x}$. The interpretation of the recursivity is now plain : the average increase of concentration resulting from a merging of two entries is linear, the simplest form of increase we can find.

Theorem 1. *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ satisfies axioms 2-4 if and only if $C(\mathbf{x}) = \text{HH}(\mathbf{x})$.*

Proof. (\Rightarrow)

Note that we can rewrite

$$\text{HH}(\mathbf{x}) = \frac{\sum_{i=1}^n (x_i)^2}{(\sum_{i=1}^n x_i)^2}$$

to see directly that

$$\text{HH}(k\mathbf{x}) = \frac{\sum_{i=1}^n (kx_i)^2}{(\sum_{i=1}^n kx_i)^2} = \frac{k^2 \sum_{i=1}^n x_i^2}{k^2 (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n x_i^2}{(\sum_{i=1}^n x_i)^2}$$

for any scalar $k > 0$, so that HH satisfies axiom 1, and thus axiom 2. Let $\mathbf{x} = \mathbf{1}a$ with $a \in \mathbb{R}_+$. Then we have

$$\text{HH}(\mathbf{x}) = \frac{\sum_{i=1}^n a^2}{(\sum_{i=1}^n a)^2} = \frac{na^2}{n^2 a^2} = \frac{1}{n},$$

satisfying the axiom 3.

Now, for any $\mathbf{x} \in \mathbb{R}_+^n$,

$$\text{HH}(\mathbf{x}) = \sum_{i=1}^n s_i^2 = s_1^2 + s_2^2 + s_3^2 + \cdots + s_n^2 = (s_1 + s_2)^2 + s_3^2 + \cdots + s_n^2 - 2s_1 s_2,$$

satisfying the axiom 4.

(\Leftarrow)

For any $\mathbf{x} \in \mathbb{R}_+^n$, by axiom 2,

$$C(\mathbf{x}) = C(\mathbf{s}).$$

By axioms 3, for $x \in \mathbb{R}_+$,

$$C(x) = 1.$$

Let $S_i = \sum_{j=1}^i s_j$ and $\dot{S}_i = S_{i-1} s_i$ for all $i = 1, \dots, n-1$.

By successive applications of axiom 4,

$$\begin{aligned} C(s_1, \dots, s_n) &= C(s_1 + s_2, s_3, \dots, s_n) - 2s_1 s_2 \\ &= C(S_2, s_3, \dots, s_n) - 2\dot{S}_2 \\ &= C(S_2 + s_3, s_4, \dots, s_n) - 2\dot{S}_2 - 2S_2 s_3 \\ &= C(S_3, s_4, \dots, s_n) - 2\dot{S}_2 - 2\dot{S}_3 \\ &\vdots \\ &= C(S_{n-1}, s_n) - 2 \left(\sum_{i=2}^{n-1} \dot{S}_i \right) \\ &= C(S_{n-1}, 1 - S_{n-1}) - 2 \left(\sum_{i=2}^{n-1} \dot{S}_i \right) \\ &= C(S_{n-1} + 1 - S_{n-1}) - 2 \left(\sum_{i=2}^n \dot{S}_i \right) \\ &= 1 - 2 \left(\sum_{i=2}^n \dot{S}_i \right) \\ &= \sum_{i=1}^n s_i^2 = \text{HH}(\mathbf{x}) \end{aligned}$$

by noting that $s_n = 1 - \sum_{i=1}^{n-1} s_i = 1 - S_{n-1}$.

□

The axiomatization of the ENP can be done straightforward by some simple adaptation of the axioms, since the ENP can be expressed as the inverse of HH.

Axiom 5 (Reflexivity*). *For all $n \in \mathbb{N}$, for all $a \in \mathbb{R}_+$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is reflexive if and only if*

$$C(\mathbf{1}a) = n$$

with $\mathbf{1} \in \mathbb{R}_+^n = 1, \dots, 1$.

Axiom 6 (Recursivity*). *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is recursive if and only if*

$$C(x_1, x_2, x_3, \dots, x_n) = C(x_1 + x_2, x_3, \dots, x_n) - \frac{2\bar{x}^2}{x_1 x_2}.$$

Theorem 2. *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$, a function $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ satisfies axioms 2, 5 and 6 if and only if $C(\mathbf{x}) = \text{ENP}(\mathbf{x})$.*

Other desirable properties have been proposed in the literature for the effective number of parties or HH, but they appear to be logically deducible from the conjunction of the ones elevated to the rank of axioms. One of them is zero shares independence : if a party has zero shares in a party-system, it should not influence the fragmentation/concentration.

Definition (Zero shares independence). *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$ a function $C : \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ is independent of zero entries in \mathbf{x} , if and only if*

$$C(\mathbf{x}, 0) = C(\mathbf{x}).$$

This property can be deduced from homogeneity and recursivity :

$$\begin{aligned} C(x_1, x_2, x_3, \dots, x_n, 0) &= C(x_1 + x_2, x_3, \dots, x_n, 0) - \frac{2x_1 x_2}{\bar{x}^2} \\ &= C(x_1 + x_2 + x_3, \dots, x_n, 0) - 2 \frac{x_1 x_2 - (x_1 + x_2)x_3}{\bar{x}^2} \\ &= \dots \\ &= 1 - 2 \left(\sum_{i=2}^n \dot{S}_i \right) - 2 \left(\sum_{i=1}^n x_i \right) 0 \\ &= \sum_{i=1}^n s_i^2 = C(x_1, x_2, x_3, \dots, x_n). \end{aligned}$$

Another desirable property also states that merging of two entries should not reduce concentration. This can also be satisfied thanks to the

recursivity axiom. To deepen our intuitive understanding of the functioning and interpretation of the recursivity axiom, we present the next corollary of Theorem 1.

Corollary 1. *For all $n \in \mathbb{N}$, for all $\mathbf{x} \in \mathbb{R}_+^n$ such that $x_1 = \frac{\bar{x}}{2}$, let $C : \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ be a concentration index satisfying axioms 2,5 and 6, then*

$$C(x_1 + x_2, x_3, \dots, x_n) - C(x_1, x_2, x_3, \dots, x_n) = s_2$$

that is the increase in concentration resulting from the merging of x_1 and x_2 is exactly $s_2 = \frac{x_2}{\bar{x}}$.

Let $\mathbf{s} \in [0, 1]^n$ be a seat shares distribution such that $s_1 = 50\%$, then, by the recursivity axiom we know that $\text{HH}(0.5 + s_2, s_3, \dots, s_n) - \text{HH}(\mathbf{s}) = s_2$. In particular, if $\mathbf{s} = (0.5, 0.5)$, by the reflexivity axiom $\text{HH}(\mathbf{s}) = 1/2$. By the reflexivity axiom we also have $\text{HH}(s_1 + s_2 = 1) = 1$, and thus $\text{HH}(1) - \text{HH}(\mathbf{s}) = 1/2$, as desired.

Conclusion

Various authors have proposed amelioration of the Laakso-Taagepera effective number of parties in order to measure the fragmentation without taking the time to address the fundamental question of defining what is the fragmentation of a party-system? Since the measurement of fragmentation is now intimately linked to the ENP, even exclusively, we propose a statistical interpretation of the ENP as consistent definition of party-system fragmentation. By putting the class-size paradox of Feld and Grofman [2007] in a more general statistical framework, we can tell a nice story in which fits perfectly the measurement of fragmentation and draw thereby a new interpretation by analogy. This new interpretation is the following. For a given party-system, the ENP is in expectation, the number of parties having the same size as the party owning a seat drawn at random with probability equals to 1 over the total number of seats. The inverse of the ENP, known as the Hirschman-Herfindahl concentration index, is the expected size of the party owning a seat drawn at random with the same probability. We then show how the ENP is fully characterized by the properties of zero-homogeneity, reflexivity and a particular recursion equation.

References

J. Aczel and Z. Daroczy. *On measures of information and their characterizations*. Academic Press, New York, 1975. ISBN 0120437600.

- M.A. Adelman. Comment on the "h" concentration measure as a numbers-equivalent. *The Review of economics and statistics*, 51(1):99–101, 1969.
- C. Blackorby, D. Donaldson, and J. Weymark. A normative approach to industrial-performance evaluation and concentration indices. *European Economic Review*, 19(1):89–121, 1982.
- A. Blau. The effective number of parties at four scales: Votes, seats, legislative power and cabinet power. *Party Politics*, 14(2):167–187, Mar 2008. doi: 10.1177/1354068807085888.
- S. Chakravarty and W. Eichorn. An axiomatic characterization of a generalized index of concentration. *The Journal of Productivity Analysis*, 2: 103–112, 1991.
- D.R. Cox and P.A.W. Lewis. *The statistical analysis of series of events*. Methuen, London, 1966.
- R.J. Dalton. The quantity and the quality of party systems. *Comparative Political Studies*, 41(7):899, 2008.
- S. Davies. Choosing between concentration indices: The iso-concentration curve. *Economica*, 46:67–76, 1979.
- P. Dumont and J.F. Caulier. The effective number of relevant parties : How voting power improves the Laakso-Taagepera's index. *Cahier du CEREC*, 7, 2003.
- P. Dunleavy and F. Boucek. Constructing the number of parties. *Party Politics*, 9(3):291, 2003.
- M. Duverger. *Political Parties: Their Organization and Activity in the Modern State*. London: Methuen; New York: Wiley, 1959.
- B Ebanks, P Sahoo, and W Sander. *Characterizations of information measures*. World Scientific Publishing, 1998.
- D.J. Elkins. The measurement of party competition. *The American Political Science Review*, 68(2):682–700, 1974.
- S.L. Feld. Why your friends have more friends than you do. *The American Journal of Sociology*, 96(6):pp. 1464–1477, 1991. ISSN 00029602. URL <http://www.jstor.org/stable/2781907>.
- S.L. Feld and B. Grofman. Variation in class size, the class size paradox, and some consequences for students. *Research in Higher Education*, 6:215–222, 1977. ISSN 0361-0365. URL <http://dx.doi.org/10.1007/BF00991287>.

- S.L. Feld and B. Grofman. The laakso-taagepera index in a mean and variance framework. *Journal of Theoretical Politics*, 19(1):101, 2007.
- S.L. Feld and B. Grofman. Puzzles and paradoxes involving averages: An intuitive approach. In *Collective Decision Making*, volume 43 of *Theory and Decision Library C*, pages 137–150. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-02865-6.
- W. Feller. *An introduction to probability theory and its applications Vol II*. John Wiley & Sons, 1971.
- R.A. Fischer. The effects of methods ascertainment upon the estimation of frequencies. *Annals of eugenics*, 6:13–25, 1934.
- G.V. Golosov. The effective number of parties: A new approach. *Party Politics*, 16(2):171–192, Mar 2010. doi: 10.1177/1354068809339538.
- L. Hannah and J.A. Kay. *Concentration in modern industries: Theory, measurement and the U.K. experience*. Macmillan, London, 1977.
- L. Kish. Sampling organizations and groups of unequal sizes. *American Sociological Review*, 30(4):564–572, 1965.
- R. Kline. How we count counts: The empirical effects of using coalitional potential to measure the effective number of parties. *Electoral Studies*, 28(2):261–269, May 2009. doi: 10.1016/j.electstud.2009.01.008. URL <http://dx.doi.org/10.1016/j.electstud.2009.01.008>.
- M. Laakso and R. Taagepera. Effective number of parties: A measure with application to west europe. *Comparative Political Studies*, 12(1):3–27, 1979.
- A. Lann and R. Falk. A closer look at a relatively neglected mean. *Teaching Statistics*, 27(3):76–80, 2005.
- A. Lijphart. *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945-1990*. Oxford: Oxford University Press, 1994.
- J. Molinar. Counting the Number of Parties: An Alternative Index. *American Political Science Review*, 85:1383–91, 1991.
- G.P. Patil, C.R. Rao, and M. Zelen. *Encyclopedia of statistics*, chapter Weighed distributions, pages 565–571. New-York : Wiley, 1988.
- R. Ricciuti. Political fragmentation and fiscal outcomes. *Public choice*, 118(3):365–388, 2004.
- S. Ross. *Introduction to probability models*. Academic Press, New York, 1972.

- R.L. Scheaffer. Size-biased sampling. *Technometrics*, 14(3):pp. 635–644, 1972. ISSN 00401706. URL <http://www.jstor.org/stable/1267291>.
- W.E. Stein and R. Dattero. Sampling bias and the inspection paradox. *Mathematics Magazine*, 58(2):96–99, 1985.
- R. Taagepera. *Predicting party sizes: the logic of simple electoral systems*. Oxford University Press, Jan 2007.
- R. Taagepera and M.S. Shugart. *Seats and Votes: The Effects and Determinants of Electoral Systems*. Yale University Press, New Haven, Conn., 1989.
- M. Zelen. Length-biased sampling and biomedical problems. In *Biometric Society Meeting, Dallas, Texas*, 1972.

Appendix

In this appendix, we reproduce some entries that can be found in Patil et al. [1988]. Suppose X is a nonnegative observable random variable with PDF $f(x)$. Suppose a realization x of X under $f(x)$ enters the investigator's record with probability proportional to $w(x, \beta)$, so that $Pr(Recording|X = x) = w(x, \beta)$.

Here the recording (weight) function $w(x, \beta)$ is a nonnegative function with parameter β representing the recording (sighting) mechanism. Clearly, the recorded x is not an observation on X , but on the random variable X^w having PDF

$$f^w(x, \beta) = \frac{w(x, \beta)f(x)}{\omega} \quad (9)$$

where $\omega = E[w(X, \beta)]$ is the normalizing factor, making the total probability equal to unity. The random variable X^w is called the weighted version of X , and its distribution in relation to that of X is called the weighted distribution with weight function w . An important weighted distribution corresponds to $w(x, \beta) = x$, in which case, $X^w = X^*$ is called the size-biased version of X . The distribution of X^* is called the size-biased distribution of PDF

$$f^*(x) = \frac{xf(x)}{\mu} \quad (10)$$

where $\mu = E[X]$. The PDF f^* is called the length-biased or size-biased version of f , and the corresponding observational mechanism is called length- or size-biased sampling.

Some interesting results concerning size-biased sampling : (Zelen [1972]). The expected value of the size- biased version X^* is $E[X^*] = \mu[1 + \sigma^2/\mu^2]$, where $E[X] = \mu$ and $V(X) = \sigma^2$. Furthermore, the harmonic mean of

X^* is equal to the mean of the original random variable X when it is positive, i.e., $E[1/X^*] = 1/\mu$. Another way of expressing these results is that $E[X^*]E[1/X^*] = 1 + \sigma^2/\mu^2$.